CONTINUOUS TIME MARKOV PROCESSES (CTMP)

In this chapter we consider Markov Process (MP), for which

$$\mathbb{P}(\{\text{future}\} \mid \{\text{exact present}\} \;\&\; \{\text{past}\}) = \mathbb{P}(\{\text{future}\} \mid \{\text{exact present}\}),$$

i.e.

$$\mathbb{P}(\{X_s,\ s>t\}\in A \mid X_t = x,\ \{X_u,\ u<t\}\in B) = \mathbb{P}(\{X_s,\ s>t\}\in A \mid X_t = x).$$

this holds true $\forall t$, any state x in some S = State space, and appropriate A & B (ie. belonging to some σ-algebra)

↳ We are interested in cases where $S = \{0, 1, 2, \dots\}$.

[These processes are referred to as CTMP, or Pure Jump Markov Processes, or Continuous time Markov Chains]

↳ The evolution of such processes is governed by its init. distrib. and transition probabilities $P(X_{s+t} = k \mid X_s = j)$.

If these transition proba do not depend on s, the MP is called HOMOGENEOUS, and

$$P(X_{s+t} = k \mid X_s = j) = P(X_t = k \mid X_0 = j) \equiv p_{jk}^{(t)}.$$

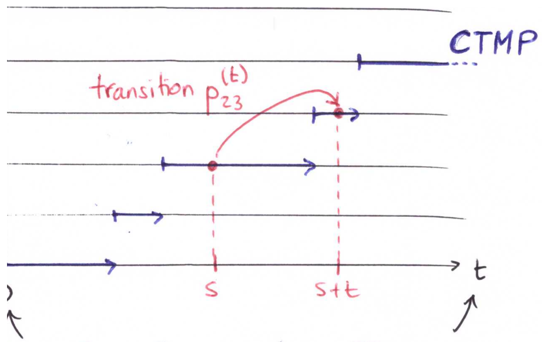
— We will only deal with such processes —

Transition probabilities satisfy the CHAPMAN-KOLMOGOROV eq (CK):

$$\boxed{P^{(t+s)} = P^{(t)} P^{(s)} \quad s,t>0}$$

where $P^{(t)} = \left(p_{jk}^{(t)}\right)$

Just like for discrete time MC: see page 6, Chapter on Markov Chains



State Space S is discrete — Time index is continuous

The FDDs can be easily obtained. let $p_j = P(X_0 = j)$

$$P(X_{t_1} = k_1, \dots, X_{t_n} = k_n) = \sum_j p_j\, p_{jk_1}^{(t_1)} \times \dots \times p_{k_{n-1}k_n}^{(t_n - t_{n-1})}$$

Goal: Find / Analyse the behaviour of the transition probabilities.

How: (i) For discrete time MC, recall that the CHAPMAN-KOLM. equation implies that at time n, $P^{(n)} = P^n$.

$P^{(n)}$ has entries $p_{jk}^{(n)}$, defined by $p_{jk}^{(n)} = P(X_n = k \mid X_0 = j)$ = n-step transition proba.

P^n: n-th power of the transition matrix.

In this case, the transition matrix P completely determines the n-step transition probabilities (and limiting behaviour, i.e. expression of the stationary distribution, provided it exists).

(ii) For continuous time MC, the time index is continuous $\Rightarrow$ we make use of differential equations to describe/analyse $P^{(t)}$. Note that

$$P^{(t+h)} - P^{(t)} \underset{CK}{=} P^{(t)}(P^{(h)} - I) = (P^{(h)} - I)P^{(t)}$$

for $h>0$; where $I = \mathrm{diag}(1, 1, \dots)$ = Identity Matrix

What's next? Well, we would like to divide both sides by h, and let $h \to 0$. But we can do this under additional technical conditions only.

↳ Indeed, we need to assume that the SP $\{X_t\}$ is STOCHASTICALLY CONTINUOUS: $X_{t+h} \to X_t$ in probability, as $h \to 0$.

"no jump" in the time interval $[t, t+h]$. For CTMP, it is enough to ensure that $P^{(h)} \longrightarrow P^{(0)} \equiv I$, as $h \to 0$ (starting somewhere at time 0, you remain there w.p 1 at time h, as $h \to 0$: again, no jump!)

$$\Rightarrow h^{-1}\left(P^{(t+h)} - P^{(t)}\right) = P^{(t)}\left(\frac{P^{(h)} - I}{h}\right) = \left(\frac{P^{(h)} - I}{h}\right) P^{(t)}$$

let $h \to 0$:

$$(*) \quad \boxed{\frac{d}{dt} P^{(t)} = P^{(t)} A = A\, P^{(t)}}$$

The derivative of the matrix is understood component-wise; i.e. with entries $\frac{d}{dt} P_{jk}^{(t)}$.

$$A \equiv (a_{jk}) := \frac{d}{dt} P^{(t)}\Big|_{t=0} := \lim_{h \to 0} \frac{P^{(h)} - I}{h}$$

is called the GENERATOR of the MP $\{X_t\}$.

Hey, what would happen if $P^{(h)} \not\to I$ as $h \to 0$?

Remarks: (i) The rows of $P^{(h)}$ sum to one $\Rightarrow$ The rows of $P^{(h)} - I$ sum to zero $\Rightarrow$ The rows of the generator sum to zero, i.e. $\sum_k a_{jk} = 0$.

(ii) System $(*)$ is a system of linear differential eq. with constant coefficients. Recall that for a univariate diff. funtion $f(t)$ satisfying $f'(t) = a\, f(t)$, it is of the form $f(t) = f(0)\, e^{at}$. It turns out that the system $(*)$ has a unique solution of a similar form, given by

$$\boxed{P^{(t)} = e^{tA} \equiv \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n, \quad t \geq 0} \leftarrow \text{Not a practical expression.}$$

With $P^{(0)} = I$, we set $A^0 = I$.

$\Rightarrow$ The generator A completely determines the transition probabilities $P_{jk}^{(t)}$ $\equiv$ just like the (one-step) transition matrix determines the n-step transition proba in discrete-time.

- Conditions for ergodicity [existence of a limiting distribution indpt of the initial state] are similar to those for MCs [and simpler, since we do not have such thing as periodicity]: if there is only one class of essential states, and $\exists\, j_0$ such that, starting at an arbitrary state j, the MP eventually visits j_0 w.p. 1, and the recurrence time to j_0 has a finite mean, then the limits

$$\lim_{t \to \infty} P_{jk}^{(t)} = \pi_k, \qquad k \in S$$

exist, independently of j: $\lim_{t \to \infty} P^{(t)} = \Pi$

all have identical rows $\pi = (\pi_k)$

↳ The distribution π is stationary for the process $\{X_t\}$:

$$\forall t \geq 0, \quad \pi = \pi P^{(t)}$$

Since $\pi = \lim_{s \to \infty} p P^{(s)} = \lim_{s \to \infty} p P^{(s-t)} P^{(t)} = \pi P^{(t)}$

p: initial distribution

Compare with expression on page 24 in the Chapter on Markov Chains.

Taking derivatives of this expression, and setting $t=0$ yields:

$$0 = \pi \frac{dP^{(t)}}{dt} = \pi A P^{(t)} \Rightarrow \boxed{\pi A = 0} \leftarrow \text{knowledge of } A \text{ is important}$$

Remark: Entries of A have a nice interpretation. Indeed,

- Recalling the definition of A: $P^{(h)} \approx P^{(0)} + hA$ for small h. Thus:

$$\mathbb{P}(X_{t+h} = k \mid X_t = j) = p^{(h)}_{jk} \approx (I + hA)_{jk} = \begin{cases} h\, a_{jk} & \text{if } j \neq k \\ 1 + h\, a_{jj} & \text{if } j = k \end{cases}$$

Since these quantities represent probabilities, one must have that $a_{jk} \geq 0$ for $j \neq k$, and $a_{jj} \leq 0$

→ $\boxed{a_{jk} = \text{rate of transition } j \mapsto k}$

- Also,

$$\mathbb{P}(X_{t+h} \neq j \mid X_t = j) = \sum_{k \neq j} \mathbb{P}(X_{t+h} = k \mid X_t = j) \approx h \sum_{k \neq j} a_{jk} = -h\, a_{jj}$$

since the rows of A sum to 0

In other words, we have that

$$\frac{\mathbb{P}(X_{t+h} \neq j \mid X_t = j)}{h} \longrightarrow -a_{jj} \ (= |a_{jj}|), \text{ as } h \to 0.$$

This quantity corresponds to a hazard function (cf page 2 in the Chapter 'Poisson Process': the hazard function characterizes the 'instantaneous probability' of occurence of an event). The limiting hazard function is constant here; corresponding to an exponentially distributed 'holding time':

→ $\boxed{\text{The MP stays at state } j \text{ for a rdm time} \sim \mathrm{Exp}(-a_{jj})}$

- Next, given that a transition from j occurred, the proba that the new value of the process is k is given by $\frac{a_{jk}}{|a_{jj}|}$.

Indeed, for $k \neq j$, and small h,

→ $\boxed{\mathbb{P}(X_{t+h} = k \mid X_{t+h} \neq j, X_t = j)}$

$$= \frac{\mathbb{P}(X_{t+h} = k, X_{t+h} \neq j, X_t = j)}{\mathbb{P}(X_{t+h} \neq j, X_t = j)}$$

$$= \frac{\mathbb{P}(X_{t+h} = k, X_t = j)}{\mathbb{P}(X_{t+h} \neq j, X_t = j)}$$

$$= \frac{\mathbb{P}(X_{t+h} = k, X_t = j)}{\mathbb{P}(X_t = j)} \times \frac{\mathbb{P}(X_t = j)}{\mathbb{P}(X_{t+h} \neq j, X_t = j)}$$

$$= \frac{\mathbb{P}(X_{t+h} = k \mid X_t = j)}{\mathbb{P}(X_{t+h} \neq j \mid X_t = j)}$$

$$\approx \frac{h\, a_{jk}}{h\, |a_{jj}|} = \boxed{\frac{a_{jk}}{|a_{jj}|}}.$$

- **Summary:** Given a CTMP $\{X_t\}$ with generator A, and initial distribution $p = (p_j)$; $p_j = \mathbb{P}(X_0 = j)$, the process evolves as follows: starting at some $X_0 = j$, the MP stays there for a random time $\sim \mathrm{Exp}(-a_{jj})$ then it jumps to another state, and this new state will be $k \neq j$ with proba $\frac{a_{jk}}{|a_{jj}|}$.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & & \\ a_{21} & a_{22} & \cdots & & \\ \vdots & & & & \\ a_{j1} & a_{j2} & \cdots & \boxed{a_{jj}} \cdots & \boxed{a_{jk}} \cdots \\ \vdots & \vdots & & & \end{pmatrix}$$

Example: Poisson Process $\{N_t\}$ with rate λ. (7)

↘ By definition of a PoP(λ), we have that for $j=0,1,2,\dots, t\geq 0$,

$$\begin{cases} P^{(t)}_{j,j+k} = \dfrac{(\lambda t)^k}{k!} e^{-\lambda t} , & k=0,1,2,\dots \\ P^{(t)}_{j,j+k} = 0 , & k<0 \end{cases}$$

↑ derivatives of these quantities evaluated at $t=0$:

- $\dfrac{d}{dt} P^{(t)}_{jj}\Big|_{t=0} = \dfrac{d}{dt} e^{-\lambda t}\Big|_{t=0} = -\lambda e^{-\lambda t}\Big|_{t=0} = -\lambda$
- $\dfrac{d}{dt} P^{(t)}_{j,j+1}\Big|_{t=0} = \dfrac{d}{dt} \lambda t e^{-\lambda t}\Big|_{t=0} = \lambda e^{-\lambda t} - \lambda^2 t e^{-\lambda t}\Big|_{t=0} = \lambda$
- $\dfrac{d}{dt} P^{(t)}_{j,j+k}\Big|_{t=0} = 0$, for $k>1$.

↘ The generator of a PoP(λ) is thus $A = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \dots \\ 0 & -\lambda & \lambda & 0 & \dots \\ 0 & 0 & -\lambda & \lambda & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$,

And indeed, N_t remains at a given state for a random time $\sim \mathrm{Exp}(\lambda)$, and given a transition occurred, it must be one step up.
Also, recall the third definition of a PoP(λ) on page 10 in the Chapter 'Poisson Process': in addition to indpt & stationary increments, one has $P(N_h = 1 \mid N_0 = 0) = \lambda h (1 + o(1))$. And indeed, in view of ~~the~~ the interpretation of the entries of the generator on page 5, we have that $P(N_h = j+1 \mid N_0 = j) \approx a_{j,j+1} h = \lambda h$. Good.

I. BIRTH & DEATH PROCESSES. (8)

. A continuous time Markov process, with state space $S=\{0,1,2,\dots\}$ is called a Birth & Death Process (B&D) if, for its generator $A=(a_{jk})$, one has $a_{jk}=0$ if $|j-k|>1$. In other words, if the process is in state k at some time t, it can only jump to state $(k+1)$ or $(k-1)$ from there (or sits in state k for some time).

. The values • $\lambda_j = a_{j,j+1}$ are called BIRTH rates
• $\mu_j = a_{j,j-1}$ are called DEATH rates ($\mu_0 = 0$).

$$\Rightarrow \quad A = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 & \dots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & 0 & \dots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & \dots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \lambda_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

× Recall that the rows of the generator of a CTMP sum to zero.

× As $h\to 0$, $\begin{cases} P(X_{t+h} = k+1 \mid X_t = k) = \lambda_k h + o(h) \\ P(X_{t+h} = k-1 \mid X_t = k) = \mu_k h + o(h) \\ P(X_{t+h} = k \mid X_t = k) = 1 - (\lambda_k + \mu_k) h + o(h) \end{cases}$

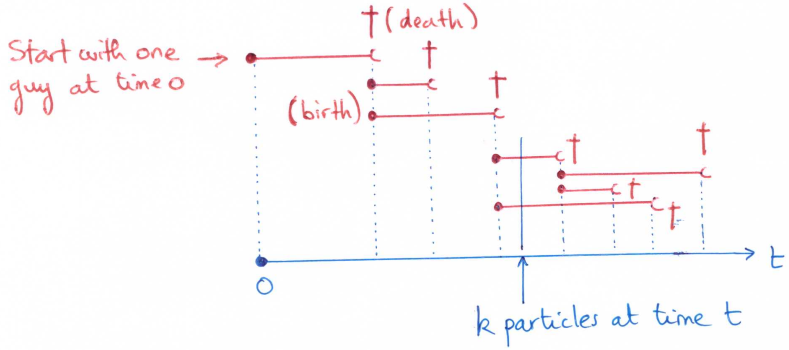
[transitions are only possible to neighbouring states]

- A B&D process sits at state k for a random time $\sim \mathrm{Exp}(\lambda_k + \mu_k)$, and then jumps to state $(k+1)$ w.p. $\lambda_k / \lambda_k + \mu_k$, or to state $(k-1)$ w.p. $\mu_k / \lambda_k + \mu_k$.

• Note that the PoP(λ) is a special case of a B&D process: it is a <u>pure birth process</u> with rates $\lambda_k = \lambda$, and $\mu_k = 0$, $k \geqslant 0$.

<u>Example</u> = Binary continuous time branching process.

→ Each particle lives for a random time $\sim \text{Exp}(\lambda)$, $\lambda > 0$, and then either splits into two new particles (w.p. p), or just dies (w.p. $q \equiv 1-p$). All particles evolve independently.



→ Given $X_t = k$, time until the next death (i.e. when a transition occurs) is the minimum of k independent exponential r.v's.
⇒ it has an $\text{Exp}(k\lambda)$ distribution.

When an event occurs, it corresponds to a transition from state k to state $(k+1)$ w.p. p, or to state $(k-1)$ w.p. q.

birth ↓ Transition λ_k must be such that

$$\frac{\lambda_k}{k\lambda} = p \Rightarrow \lambda_k = pk\lambda$$

death ↓ Similarly, we get $\mu_k = q\lambda k$

see discussion bottom of page 6

diagonal element

$$\Rightarrow \quad A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ q\lambda & -\lambda & p\lambda & 0 & 0 & \cdots \\ 0 & 2q\lambda & -2\lambda & 2p\lambda & 0 & \cdots \\ 0 & 0 & 3q\lambda & -3\lambda & 3p\lambda & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

First row: 0 is an absorbing state.

→ Note that there exists a trivial stationary distribution $\pi = (1, 0, 0, \ldots)$, which is solution to $\pi A = 0$, with $\sum \pi_j = 1$.

<u>When $p \leqslant q$</u>, the process is ergodic with that stationary distrib.: the process eventually becomes extinct.

<u>When $p > q$</u>, there is no ergodicity. With a probability depending on the initial state, the process becomes extinct, while on the complementary event, one has explosion: $X_t \to \infty$ exponentially fast as $t \to \infty$.

<u>Remark</u>: • Consider

$$\varphi(t,z) = E\left(z^{X_t} \mid X_0 = 1\right) = \sum_{k \geqslant 0} p_{1k}^{(t)} z^k = \text{generating function of } X_t \text{ given } X_0 = 1.$$

Note that $E\left(z^{X_t} \mid X_0 = m\right) = \left[\varphi(t,z)\right]^m$ since the process X_t can then be thought of the sum of n independent realizations of the branching process starting with one particle.

• Making use of $\dfrac{dp^{(t)}}{dt} = A p^{(t)}$, with

$$p^{(t)} = \begin{pmatrix} p_{00}(t) & p_{01}(t) & \cdots & p_{0k}(t) & \cdots \\ p_{10}(t) & p_{11}(t) & \cdots & p_{1k}(t) & \cdots \\ p_{j0}(t) & p_{j1}(t) & \cdots & p_{jk}(t) & \cdots \\ \vdots & \vdots & & \vdots & \end{pmatrix}$$

$$A p^{(t)} = \begin{array}{c} \\ \left(\begin{array}{ccc} & k+1 & \\ & \vdots & \\ \cdots & \sum_{m \geqslant 0} a_{jm} p_{mk}(t) & \cdots \\ & \vdots & \end{array} \right) \end{array} \begin{array}{c} \\ \\ \leftarrow j+1 \\ \\ \end{array}$$

we get

$$\frac{\partial \varphi(t,z)}{\partial t} = \sum_{k\geq 0} \frac{\partial p_{1k}^{(t)}}{\partial t} z^k$$

$$= \sum_{k\geq 0}\left(\sum_{m\geq 0} a_{1m}\, p_{mk}^{(t)}\right) z^k$$

$$= \sum_{m\geq 0} a_{1m}\left(\sum_{k\geq 0} p_{mk}^{(t)} z^k\right)$$

$$= \sum_{m\geq 0} a_{1m}\, \varphi^m(t,z)$$

$$\boxed{\frac{\partial \varphi(t,z)}{\partial t} = \lambda\left(q - \varphi(t,z) + p\,\varphi^2(t,z)\right)}$$

for a binary continuous time branching process.

↳ By analyzing this differential equation, one can extract a lot of information about the behaviour of the branching process.

- Alternatively, one can make use of $\frac{\partial P^{(t)}}{\partial t} = P^{(t)} A$, which yields

$$\frac{\partial \varphi(t,z)}{\partial t} = \sum_{m\geq 0} P_{1m}^{(t)}\left(\sum_{k\geq 0} a_{mk} z^k\right)$$

$$= \sum_{m\geq 1} P_{1m}^{(t)}\Big(\underbrace{a_{m,m-1}}_{= mq\lambda} z^{m-1} + \underbrace{a_{m,m}}_{= -m\lambda} z^m + \underbrace{a_{m,m+1}}_{= mp\lambda} z^{m+1}\Big)$$

$$= \lambda q \sum_{m\geq 1} P_{1m}^{(t)} m z^{m-1} - \lambda z \sum_{m\geq 1} P_{1m}^{(t)} m z^{m-1} + \lambda p z^2 \sum_{m\geq 1} P_{1m}^{(t)} m z^{m-1}$$

$$\boxed{\frac{\partial \varphi(t,z)}{\partial t} = \frac{\partial \varphi(t,z)}{\partial z}\left\{\lambda q - \lambda z + \lambda p z^2\right\}}$$

For a B&D process, one can find the unconditional proba $p_j(t) = P(X_t = j)$ for a given initial distribution p by solving the system of diff. equations

$$\boxed{\begin{aligned} &\frac{d}{dt} p(t) = p(t)\, A \\ &p(t) = (p_0(t), p_1(t), \dots) \end{aligned}}$$

Since $p(t) = p\, P^{(t)}$, and $\frac{d}{dt} P^{(t)} = P^{(t)} A$ (see page 3)

The system has the form

$$\begin{cases} p_0'(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t) \\ p_k'(t) = \lambda_{k-1} p_{k-1}(t) - (\lambda_k + \mu_k) p_k(t) + \mu_{k+1} p_{k+1}(t), \quad k \geq 1 \end{cases}$$

System governing the redistribution of mass as time passes.

In the steady state, $p_k(t) \to \pi_k$, $k \geq 0$, and we obtain

$$\begin{cases} 0 = -\lambda_0 \pi_0 + \mu_1 \pi_1 \\ 0 = \lambda_{k-1}\pi_{k-1} - (\lambda_k + \mu_k)\pi_k + \mu_{k+1}\pi_{k+1} \end{cases}$$

Nothing else but $0 = \pi A$ (see page 4)

This system can be solved recursively: the k-th equation can be rewritten as

$$-\lambda_{k-1}\pi_{k-1} + \mu_k \pi_k = \underbrace{-\lambda_k \pi_k + \mu_{k+1}\pi_{k+1}}, \quad k \geq 1$$

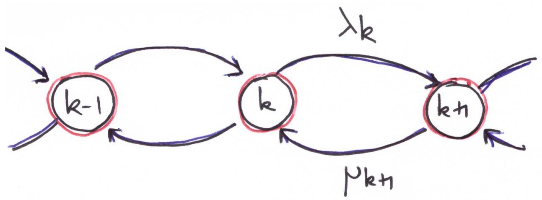
This term is the same as on the LHS, replacing k by $k+1$.
⇒ This term must be constant with k, and equal to zero since from the first equation one have $-\lambda_0\pi_0 + \mu_1\pi_1 = 0$.

Thus ←

Stationary distribution.

$$\boxed{\lambda_k \pi_k = \mu_{k+1}\pi_{k+1}, \quad k \geq 0}$$

B&D



λ_k = rate at which the mass is transferred from k to $(k+1)$

μ_{k+1} = —— " —— " —— " —— $(k+1)$ to k.

$\Rightarrow$ Equation $\lambda_k \pi_k = \mu_{k+1} \pi_{k+1}$ means that there is some sort of 'balance' in the way the mass is moved around at the stationary regime: what goes out, goes in.

Assuming $\lambda_j > 0$, $\mu_{j+1} > 0$, $j = 0, 1, \ldots$ we get

$$\pi_{k+1} = \frac{\lambda_k}{\mu_{k+1}} \pi_k = K_{k+1} \pi_0 \quad , \text{with} \begin{cases} K_{k+1} = \dfrac{\lambda_0 \lambda_1 \cdots \lambda_k}{\mu_1 \mu_2 \cdots \mu_{k+1}} , & k \geq 0 \\ K_0 = 1 \end{cases}$$

The B&D is ergodic iff $\sum_{j \geq 0} K_j < \infty$. If this is the case, we have from $1 = \sum_{j \geq 0} \pi_j = \pi_0 \sum_{j \geq 0} K_j$, that the stationary distribution is

$$\boxed{\begin{aligned} \pi_0 &= \Big(\sum_{j \geq 0} K_j \Big)^{-1} \\ \pi_j &= K_j \pi_0 , \quad j \geq 1 \end{aligned}}$$